

Statistical Considerations in Biosimilar Assessment Using Biosimilarity Index

Aijing Zhang¹, Jung-Ying Tzeng¹ and Shein-Chung Chow^{2*}

¹Department of Statistics at NC State University, Raleigh, NC 27695, USA

²Department of Bioinformatics and Biostatistics at Duke University, Durham, NC 27708, USA

Abstract

When an innovative biologic product goes off patent, biopharmaceutical or biotechnological companies may file an application for regulatory approval of biosimilar products. Unlike small molecule drug products, biosimilars are not exact copies of their brand-name counterpart, and they are usually very sensitive to changes in environmental factors and have greater variabilities due to their complexity and sensitivity to variation in manufacturing processes. Facing these challenges, a biosimilarity index based on reproducibility probability is proposed to assess biosimilarity. In this article, we have demonstrated how to assess biosimilarity between the test and reference product in relative to a reference standard that is established in a study where reference product is compared with itself. Biosimilarity index approach is robust against biosimilarity criteria and has the advantage of allowing the assessment of the degree of similarity.

Keywords: Highly similar; Biosimilarity index; Reproducibility probability; Degree of similarity

Abbreviations: BE: Bioequivalence; BPCI: Biologics Price Competition and Innovation; CI: Confidence Interval; CV: Coefficient of Variation; FDA: Food and Drug Administration; MLE: Maximum Likelihood Estimator; TOST: Two One-Sided Tests; US: United States

Introduction

According to the Biologics Price Competition and Innovation (BPCI) Act passed by the United States (US) congress in 2009, a biosimilar product is defined as a biological drug product that is highly similar to the reference product notwithstanding minor differences in clinically inactive components and there are no clinically meaningful differences in terms of safety, purity, and potency.

Biological drug products contain active ingredients that are derived from or made by living cells or organisms. Unlike generic chemical drugs, biosimilar products are expected to have similar but not identical active ingredients as the innovative brand name product. Furthermore, biological products are very sensitive to small changes at various stages of the manufacturing process and environmental factors such as light and temperature. Therefore, biosimilars have greater inherent variability than chemical drugs. Current regulations for the assessment of bioequivalence (BE) for generic chemical drugs only concern with average bioequivalence. The main criticism against the criteria for BE studies is that they do not take variabilities of the drug products into consideration. Given the greater variabilities of the biological drugs it is recognized that current regulations and/or criteria for the assessment of BE may not be applicable for the assessment of biosimilarity between biologic products.

The BPCI Act as part of the Affordable Care Act was signed into law in March 2010 which gave the US Food and Drug Administration (FDA) the authority to approve similar biological drug products. However, currently the FDA has not put forward clear standards for biosimilar approval [1] due to the complexity of the biological drug products. Facing these challenges, a new biosimilarity index approach was proposed by Chow et al. [2] to assess biosimilarity. This approach has the advantage of being robust to the study endpoints, criteria and study designs. Thus a universal approach for biosimilarity assessment

could be implemented without well-accepted criteria by the regulatory agencies. The BPCI Act also introduced the term “highly similar” in the definition of biosimilarity, but there is little or no discussion regarding how similar is considered highly similar. The proposed biosimilarity index approach can also answer the question of “how similar is highly similar” as the index quantifies the degree of similarity.

The purpose of the paper is to illustrate how to operationalize the biosimilarity index approach, and to evaluate the performance of biosimilarity index under a crossover design using average similarity criterion.

In the next section, biosimilar index based on reproducibility probability is briefly introduced. In Section 3, the statistical properties of the biosimilarity index are discussed through simulation studies. In Section 4, an example is given to further illustrate the impact of the variability on the conclusion of biosimilarity. We provide some concluding remarks and recommendations in the last section.

Biosimilarity Index

In order to reflect the characteristics and the impact of variability on the therapeutic effect of biologic products, Chow et al. in 2011 proposed the development of an index based on the concept of the reproducibility probability to evaluate the degree of similarity between two drug products [2]. Reproducibility probability was first considered by Shao and Chow [3] to address the question of whether the observed significant result from a clinical trial is reproducible.

The biosimilarity index is evaluated as the estimated power of the

***Corresponding author:** Shein-Chung Chow, Department of Bioinformatics and Biostatistics, Duke University, Durham, NC 27708, USA, E-mail: aijing.zhang@gmail.com

Received July 19, 2013; **Accepted** August 26, 2013; **Published** September 02, 2013

Citation: Zhang A, Tzeng JY, Chow SC (2013) Statistical Considerations in Biosimilar Assessment Using Biosimilarity Index. J Bioequiv Availab 5: 209-214. doi:10.4172/jbb.1000160

Copyright: © 2013 Zhang A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

testing procedure when the alternative hypothesis is true, replacing the parameter by its estimate based on the observed data. The hypotheses of the similarity testing are often expressed as two sets of one-sided hypotheses:

$$H_{01} : \theta_L \text{ vs. } H_{a1} : \theta > \theta_L \text{ and } H_{02} : \theta \geq \theta_U \text{ vs. } H_{a2} : \theta < \theta_U \quad (1)$$

Where, θ is the study parameter chosen to assess biosimilarity; θ_L and θ_U are the biosimilarity limits, i.e., the accepted lower and upper bounds for declaring biosimilarity.

The evaluation of biosimilarity index depends on the form of the test statistics, which in turn depend on the study designs and the criteria chosen. For the 2x2 crossover design, we consider the following statistical model:

$$Y_{ijk} = \mu + S_{ik} + P_j + T_{(i,k)} + \varepsilon_{ijk} \quad (2)$$

where Y_{ijk} is the response for subject i in the k_{th} sequence at the j_{th} period, where $i=1, \dots, n_k$ indicates subject, $j=1, 2$ indicates period, $k=1, 2$ indicates sequence; μ represents the overall mean; S_{ik} represents the random effect of i_{th} subject in k_{th} sequence, assumed independently and identically distributed (i.i.d.) as $N(0, \sigma)$; P_j is the fixed period effect; $T_{(i,k)}$ represents the fixed effect of the treatment in the k_{th} sequence administered at the j_{th} period; ε_{ijk} is the within-subject random error, assumed i.i.d. as $N(0, \sigma)$. Finally

S_{ik} 's and ε_{ijk} 's assumed to be mutually independent.

When we choose average biosimilarity criterion, i.e., $\theta = \mu_T - \mu_R$, the test statistics for Equation (1) are:

$$T_L = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_L}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ and } T_U = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_U}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (3)$$

Where, \bar{Y}_T and \bar{Y}_R are the least square means for the test and reference products; they can be obtained from the sequence-by-period means: $\bar{Y}_R = \frac{1}{2}(\bar{Y}_{11} + \bar{Y}_{22})$ and $\bar{Y}_T = \frac{1}{2}(\bar{Y}_{12} + \bar{Y}_{21})$; n_1 and n_2 are the number of subjects in sequence 1 and sequence 2; $\hat{\sigma}_d^2$ is the pooled sample variance of period difference from both sequences and can be obtained from σ

$$\hat{\sigma}_d^2 = \frac{1}{n_1 + n_2 - 2} \sum_{k=1}^2 \sum_{i=1}^{n_k} (d_{ik} - \bar{d}_k)^2,$$

where $(d_{ik} = \frac{1}{2}(Y_{i2k} - Y_{i1k}), i=1, 2, \dots, n_k; k=1, 2)$. By the estimated

power approach, the biosimilarity index \hat{P}_{BI} for the 2x2 crossover study using average biosimilarity criterion can be obtained:

$$\hat{P}_{BI} = P(T_L(Y) > t_L \text{ and } T_U(Y) < t_U | \hat{\sigma}_L, \hat{\sigma}_U) \quad (4)$$

Where, T_L and T_U are the test statistics given in Equation (3). Both T_L and T_U follow non-central t-distribution, with $n_1 + n_2 - 2$ degree of freedom and non-centrality parameters δ_L and δ_U respectively; δ_L and δ_U relate to the population means, variances and biosimilarity limits; their estimate $\hat{\delta}_L$ and $\hat{\delta}_U$ can be obtained from the data using the formula

$$\hat{\delta}_L = \frac{(\bar{Y}_T - \bar{Y}_R - \theta_L)}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ and } \hat{\delta}_U = \frac{(\bar{Y}_T - \bar{Y}_R - \theta_U)}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\hat{\delta}_U = \frac{(\bar{Y}_T - \bar{Y}_R - \theta_U)}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

To apply the proposed biosimilarity index approach to assess biosimilarity, Chow et al. [2] proposed the following steps [4]:

Step 1: Assess the average biosimilarity based on a given biosimilarity criterion. The criterion could be based on mean, ratio or variability.

Step 2: Once the product passes the test for biosimilarity in Step 1, calculate biosimilarity index of Equation (4) based on the observed mean difference and standard deviation. The calculated biosimilarity index thus takes the variability and the sensitivity to heterogeneity in variances into consideration.

Step 3: We then claim highly biosimilar if the calculated 95% confidence lower bound of the biosimilarity index is larger than p_0 , a pre-specified limit on declaring highly biosimilar.

To establish p_0 , we recommend that it be based on p_{RR} , the biosimilarity index obtained in an R-R study where reference product is compared with itself. By basing p_0 on p_{RR} , the biosimilarity index approach allows us to assess the degree of similarity in relative to the reference product [5].

From the definition of the biosimilarity index and the testing steps outlined above, we can see that this approach has several advantages. First, it is robust with respect to the selected study endpoint, biosimilarity criterion and study design [4] because the biosimilarity index utilized in the second stage of testing "highly similar" is calculated using the same selected study endpoint, biosimilarity criterion and study design. Second, it takes variability into consideration for the calculation of the index, and is sensitive to the variance of the test products. Third, it allows the assessment of the degree of similarity. Other words, it provides an answer to the question of "how similar is considered similar?".

Numerical Results

The proposed biosimilarity index approach as outlined in Section 2 is implemented in simulation studies to demonstrate the statistical properties of the index. Standard 2x2 crossover study design and average biosimilarity criterion are used. The biosimilarity index is calculated as in Equation (4).

Simulation design

The study parameter θ is the mean difference between test and reference products, i.e., $\theta = \mu_T - \mu_R$, and the corresponding similarity limits are those accepted in BE studies, i.e., $(\theta_L, \theta_U) = (\log 0.8, \log 1.25) = (-0.223, 0.223)$. The true value of θ has been chosen to range from 0 to 0.15. In order to evaluate how biosimilarity indices are affected by the variabilities of the drug products, different variances for the test product are chosen: 0.2, 0.25, or 0.3 while the variance of the reference is kept at 0.3. The impact of different total sample sizes ($n=n_1+n_2$) with equal allocation is also investigated. The parameter settings for the simulation studies are summarized in Table 1.

Results

A total of 1,000 random trials are generated for each combination of the parameter specifications. Table 2 records the percentage of trials that have passed the Step 1 biosimilarity test, i.e., the probability of

Sample Size	Population parameters for R and T average		Biosimilarity criteria
	(μ_R, μ_T)	(σ_R, σ_T)	
n			(θ_L, θ_U)
40	(1, 1)	(0.30, 0.25)	(-0.223, 0.223)
50	(1, 1.025)	(0.30, 0.30)	
60	(1, 1.05)	(0.30, 0.35)	
70	(1, 1.075)		
80	(1, 1.1)		
90	(1, 1.125)		
100	(1, 1.15)		

Table 1: Parameter specification for the simulation studies.

claiming biosimilarity on the basis of the average biosimilarity criterion. As the mean difference between the test and reference products increase, the probability of claiming biosimilarity decreases. When the variance of the test product increases, the probability decreases as well. Increasing sample size can help increase the probability of claiming biosimilarity. In typical BE studies, the sample sizes range from 18 to 24. To assess biosimilarity, the sample size for the comparative nonclinical and clinical studies are expected to be larger than those chosen in BE studies, but the studies are still conducted in limited number of patients when compared with that used in the pivotal trials for the innovative drugs.

For those trials that have passed the Step 1 test, Table 3 reports the average of the p-values obtained from the Schuirmann's two one-sided tests (TOST) procedure. As the mean difference between the test and reference products increase, the p-value increases. When the variance of the test product increases, the p-value increases as well. In another word, as the mean difference and/or variance increase, for those trials where we are able to declare biosimilarity, the evidence against null

hypotheses nonetheless weakens.

The biosimilarity index, i.e., the reproducibility probability in Equation (4) is calculated as the steps outlined in Section 2. Table 4 further shows the value of p_{TR} for those trials that have passed the Step 1 test. As expected, the results show that the p_{TR} decreases as the mean difference or variance increases; and it increases as the

Next we calculate the percentage of trials that pass the "highly similar" test based on the biosimilarity index p_{TR} . For the p in Step 3 of the testing procedure proposed in Section 2, we choose it to be 0.8 PRR p_{RR} where PRR p_{RR} is assumed known and constant, set at 80%. As the mean difference between test and reference products increases or the variance of the test product increases, the percentage of passing "highly similar" test decreases. The percentage of passing increases as the sample size increases.

Note that when the mean difference between test and reference products is large, such as the difference is 0.15, the test drug could not pass the "highly similar" test even if we declared biosimilarity in Step 1. This is due to the fact that null hypothesis in Step 1 was rejected on weak evidence. Or in another word, when we want to make claims on the degree of similarity, additional information is utilized in order to quantify the similarity in comparison with the reference product.

Finally, when the test product has a larger variance than the reference product, the results show that it gets harder to conclude the same level of similarity. This demonstrates the proposed biosimilarity index approach is sensitive to the heterogeneity in variances, and can reflect the impact of variability of the biological products.

Example

As shown in the simulation studies, as coefficient of variation (CV) gets bigger, it is less likely that we are able to declare similarity

σ_T	Sample Size (n)	Mean Difference ($\mu_T - \mu_R$)						
		0%	2.50%	5%	7.50%	10%	12.50%	15%
0.25	40	96.50%	94.50%	89.40%	79.00%	64.20%	47.90%	31.90%
	50	98.90%	97.60%	93.50%	84.00%	73.50%	55.20%	36.10%
	60	99.60%	98.70%	96.00%	89.70%	78.20%	60.70%	40.80%
	70	99.90%	99.60%	97.90%	92.30%	82.30%	66.70%	46.00%
	80	99.90%	99.80%	99.30%	96.00%	86.10%	70.60%	50.80%
	90	100%	99.80%	99.40%	96.50%	88.80%	75.80%	54.00%
0.3	40	93%	90.00%	83.80%	72.40%	57.10%	41.60%	28.70%
	50	96.50%	94.50%	89.40%	80.10%	67.90%	49.50%	31.80%
	60	98.40%	97.10%	92.70%	85.30%	72.60%	54.60%	36.90%
	70	99.50%	98.90%	95.50%	88.60%	77.30%	60.60%	41.50%
	80	99.80%	99.50%	97.10%	92.50%	82.00%	64.90%	44.90%
	90	99.80%	99.60%	98.40%	94.50%	84.80%	68.50%	48.70%
0.35	40	85.50%	84.00%	76.50%	65.50%	52.00%	37.60%	24.80%
	50	92.20%	90.80%	84.10%	74.40%	61.00%	43.50%	27.40%
	60	95.80%	94.50%	88.90%	79.70%	65.90%	49.10%	32.80%
	70	98.50%	96.80%	92.00%	83.50%	71.40%	54.30%	38.00%
	80	99.60%	98.40%	95.30%	88.20%	75.00%	58.60%	40.60%
	90	99.60%	99.00%	96.50%	89.30%	79.70%	63.10%	44.10%

Table 2: Percentage of trials that passed the Step 1 test.

σ_T	Sample Size (n)	Mean Difference ($\mu_T - \mu_R$)						
		0%	2.50%	5%	7.50%	10%	12.50%	15%
0.25	40	0.0055	0.006	0.0079	0.0106	0.0135	0.0162	0.0176
	50	0.003	0.0036	0.0053	0.0076	0.0114	0.0154	0.0175
	60	0.0016	0.0021	0.0038	0.0064	0.01	0.0137	0.0172
	70	0.0008	0.0012	0.0028	0.005	0.0082	0.0122	0.0157
	80	0.0003	0.0006	0.0019	0.0041	0.0071	0.011	0.0158
	90	0.0002	0.0004	0.0014	0.0032	0.0059	0.0103	0.0142
0.3	40	0.0081	0.0086	0.0104	0.0126	0.0145	0.0162	0.0182
	50	0.0047	0.0052	0.007	0.0094	0.0135	0.0166	0.0179
	60	0.003	0.0036	0.0052	0.0082	0.0118	0.0148	0.0184
	70	0.0018	0.0025	0.0042	0.0066	0.0098	0.0134	0.0163
	80	0.0009	0.0014	0.0027	0.0055	0.0091	0.0126	0.0166
	90	0.0005	0.0009	0.0023	0.0048	0.0075	0.0111	0.0153
0.35	40	0.0106	0.0119	0.0132	0.0151	0.0169	0.0176	0.0174
	50	0.0068	0.0076	0.0089	0.0114	0.0151	0.0174	0.0171
	60	0.0047	0.0055	0.0071	0.0099	0.0131	0.0159	0.0192
	70	0.0034	0.004	0.0055	0.0079	0.0113	0.0141	0.0174
	80	0.0022	0.0025	0.0042	0.007	0.0099	0.0138	0.0175
	90	0.0012	0.0018	0.0035	0.0054	0.0091	0.0128	0.0178

Table 3: Average of the p-values for trials that passed the Step 1 test.

σ_T	Sample Size (n)	Mean Difference ($\mu_T - \mu_R$)						
		0%	2.50%	5%	7.50%	10%	12.50%	15%
0.25	40	0.873	0.864	0.839	0.808	0.779	0.752	0.736
	50	0.922	0.912	0.885	0.85	0.88	0.76	0.736
	60	0.949	0.939	0.91	0.868	0.819	0.773	0.732
	70	97%	96%	93%	89%	84%	79%	75%
	80	98%	98%	95%	91%	86%	80%	75%
	90	99%	98%	96%	93%	88%	81%	76%
0.3	40	82%	82%	80%	78%	76%	74%	73%
	50	0.885	0.876	0.852	0.819	0.775	0.746	0.731
	60	92%	91%	88%	84%	80%	76%	72%
	70	94%	93%	91%	87%	82%	78%	74%
	80	97%	96%	93%	88%	83%	79%	74%
	90	98%	97%	994%	90%	85%	80%	75%
0.35	40	0.772	0.76	0.748	0.733	0.72	0.716	0.719
	50	0.84	0.831	0.814	0.785	0.751	0.733	0.732
	60	0.878	0.868	0.846	0.812	0.776	0.744	0.71
	70	0.908	0.9	0.878	0.844	0.802	0.768	0.73
	80	0.936	0.928	0.9	0.858	0.816	0.77	0.728
	90	0.955	0.946	0.918	0.882	0.83	0.782	0.738

Table 4: Biosimilarity Index - p_{TR} .

even when there is no true mean difference. In this section, we use an example data to further illustrate the impact of the high variability on the conclusion of biosimilarity and how the biosimilarity index assesses the degree of similarity (Table 5).

In the simulation studies above, we have considered the scenario where p_{RR} is constant. This constant p_{RR} could be obtained from a separate R-R study [5]. In the following example, we set out to obtain p_{RR} concurrently as we assess the test product [6], and thus is

considered random. To obtain p_{RR} concurrently, a slight variation of the 2x2 crossover design is used. Namely, for the first sequence, subjects are treated with T at the first dosing period, and crossover to the second dosing period to receive R. However for the R treatment, subjects are again randomly split into two groups, and treated with either R_1 or R_2 . Similarly, for the second sequence, subjects will be split into two groups treated with either R_1 or R_2 respectively, then both of these groups will crossover to receive T in the second dosing period. In this case, the design essentially becomes a 4x2 crossover design (Table 6)

We generate a sample data with a total sample size of 160, i.e., 40 subjects per sequence. The means and variances of the two reference products are assumed to be the same, i.e., $\mu_{R_1} = \mu_{R_2} = \mu_R$ and $\sigma_{R_1} = \sigma_{R_2} = \sigma_R$. We further assume there is no true mean difference between test and reference product either. A CV of 30% is chosen for all reference and test products.

When R is compared with itself, biosimilarity could not always be declared; the probability of declaring biosimilarity depends on the CV when sample size is fixed. An example where similarity between different batches of R product could not be established is obtained and the sample means are given in Table 7. The observed mean difference $\bar{Y}_{R_2} - \bar{Y}_{R_1} = 0.126$ and the observed variability $\hat{\sigma}_{d_R} = 0.433$. The 90% CI of the mean difference is (0.010, 0.242), thus similarity is not declared. When similarity could not be demonstrated between reference products, it is impossible to assess whether or not the test product is similar to this reference product. Careful studies should be conducted to avoid such situation.

Under the same parameter setting, another set of data is generated. From this set of data, we are able to declare biosimilarity between R_1 and R_2 . The sample means are given in Table 8, from which we could

calculate the sample mean difference between reference Products Y $\bar{Y}_{R_2} - \bar{Y}_{R_1} = -0.017$, the variability $\hat{\sigma}_{d_R} = 0.430$; and the 90% CI of the mean difference is (-0.130, 0.095). We further calculate the mean difference between test and reference products $\bar{Y}_T - \bar{Y}_R = 0.004$ and the variability $\hat{\sigma}_d = 0.430$, the 90% CI of the mean difference is (-0.108, 0.117), that is, the similarity between T and R is also declared. From the observed mean differences and variabilities, the \hat{p}_{RR} and \hat{p}_{TR} as evaluated from Equation (4) is 0.884, and 0.894, which could be used to further assess the degree of similarity.

Conclusion

The numerical results in Section 3 have shown that as variance increases, the probability of declaring biosimilarity in the Step 1 test decreases; the biosimilarity index decreases. The biosimilarity index is calculated for the trials that have passed the Step 1 test, thus reflects the characteristics of the biological products that have already been declared biosimilarity based on average biosimilarity criterion.

For the assessment of biosimilars, we should especially be aware of the higher variability of biological drugs, and its impact on the conclusion of biosimilarity testing. The results from the numerical studies have demonstrated that the biosimilarity index approach is sensitive to variance of the products. Other methods have been proposed in the literature to assess variability in addition to the assessment of average biosimilarity. Chow et al. [7] considered an approach based on the probability-based criteria for evaluating average biosimilarity, and demonstrated that the probability-based method is more sensitive to the change of variability than the moment based method. Hsieh et al. [8] has developed the statistical methodology for comparing variabilities for the assessment of biosimilarity and examined its performance under combinations of essential parameters. The advantage of the biosimilarity index approach is that it can be

σ_T	Sample Size (n)	Mean Difference ($\mu_T - \mu_R$)						
		0%	2.50%	5%	7.50%	10%	12.50%	15%
0.25	40	43%	24%	3%	0%	0%	0%	0%
	50	93%	76%	23%	3%	0%	0%	0%
	60	100%	89%	58%	5%	0%	0%	0%
	70	100%	100%	86%	26%	2%	0%	0%
	80	100%	100%	100%	48%	2%	0%	0%
	90	100%	100%	100%	75%	11%	0%	0%
0.3	40	6%	6%	1%	0%	0%	0%	0%
	50	55%	36%	9%	0%	0%	0%	0%
	60	89%	69%	18%	5%	0%	0%	0%
	70	100%	98%	54%	7%	0%	0%	0%
	80	100%	99%	79%	19%	0%	0%	0%
	90	100%	100%	98%	29%	6%	0%	0%
0.35	40	0%	0%	0%	0%	0%	0%	0%
	50	12%	9%	2%	1%	0%	0%	0%
	60	38%	29%	2%	0%	0%	0%	0%
	70	84%	71%	19%	0%	0%	0%	0%
	80	98%	91%	45%	9%	0%	0%	0%
	90	100%	99%	62%	15%	0%	0%	0%

Table 5: Percent of claiming highly biosimilar based on biosimilarity index - p_B

Sequence	Period I	Period II
1(T R ₁)	T	R → R ₁
2(T R ₂)	T	R → R ₂
3(R ₁ T)	R → R ₁	T
4(R ₂ T)	R → R ₂	T

Table 6: 4×2 Crossover Design.

Sequence	Period I	Period II	Sequence Mean
1	1T = $\bar{Y}_{.11} = 0.920$	1 R ₁ = $\bar{Y}_{.21} = 1.010$	$\bar{Y}_{.1} = 0.965$
2	2T = $\bar{Y}_{.12} = 0.987$	2 R ₂ = $\bar{Y}_{.22} = 0.958$	$\bar{Y}_{.2} = 0.972$
3	3 R ₁ = $\bar{Y}_{.11} = 1.063$	3T = $\bar{Y}_{.21} = 0.910$	$\bar{Y}_{.3} = 0.987$
4	4 R ₂ = $\bar{Y}_{.12} = 1.016$	4T = $\bar{Y}_{.22} = 0.997$	$\bar{Y}_{.4} = 1.007$
Period mean	$\bar{Y}_{.1.} = 0.997$	$\bar{Y}_{.2.} = 0.969$	$\bar{Y}_{...} = 0.983$

Table 7: Example - Sample means from a study where biosimilarity between R is not declared.

Sequence	Period I	Period II	Sequence Mean
1	1T = $\bar{Y}_{.11} = 1.059$	1R ₁ = $\bar{Y}_{.21} = 0.986$	$\bar{Y}_{.1} = 1.022$
2	2T = $\bar{Y}_{.12} = 0.950$	2R ₂ = $\bar{Y}_{.22} = 1.018$	$\bar{Y}_{.2} = 0.984$
3	3R ₁ = $\bar{Y}_{.11} = 1.046$	3T = $\bar{Y}_{.21} = 0.986$	$\bar{Y}_{.3} = 1.016$
4	4R ₂ = $\bar{Y}_{.12} = 1.051$	4T = $\bar{Y}_{.22} = 1.098$	$\bar{Y}_{.4} = 1.075$
Period mean	$\bar{Y}_{.1.} = 1.027$	$\bar{Y}_{.2.} = 1.022$	$\bar{Y}_{...} = 1.024$

Table 8: Example-Sample means from a study where biosimilarity between R is declared.

applied to whatever criteria chosen: the criteria used in Step 1 testing are used again in Step 3 to quantify the level of similarity.

If we define $d = p_0 / p_{RR}$, then d can be used to address the degree of similarity and the question of “how similar is highly similar?”. In this article, we have set $d=0.8$, and claim it as highly similar. If we set $d=0.7$, and claim it as moderate similar, then the percentage of passing Step 3 test would be greater. Thus this factor d allows us to quantify the level of similarity in relative to the reference product. The regulatory agency would be able to consider the class of drugs and the impact of variabilities on clinical performance, and decide what is the tolerable difference, and on what level of similarity relative to the reference drug should be required of the biomiliar products.

References

1. (2012) Draft Guidance on Scientific Considerations in Demonstrating Biosimilarity to a Reference Product. Center for Drug Evaluation and Research, the US Food and Drug Administration.
2. Chow SC, Wang J, Endrenyi L, Lachenbruch PA (2013) Scientific considerations for assessing biosimilar products. Stat Med 32: 370-381.
3. Shao J, Chow SC (2002) Reproducibility probability in clinical trials. Stat Med 21: 1727-1742.
4. Chow SC, Yang LY, Starr A, Chiu ST (2013) Statistical methods for assessing interchangeability of biosimilars. Stat Med 32: 442-448.
5. Zhang AJ, Tzeng JY, Chow SC (2013) Establishment of reference standards in biosimilar studies.
6. Kang SH, Chow SC (2013) Statistical assessment of biosimilarity based on relative distance between follow-on biologics. Stat Med 32: 382-392.
7. Chow SC, Hsieh TC, Chi E, Yang J (2010) A comparison of moment-based and probability-based criteria for assessment of follow-on biologics. J Biopharm Stat 20: 31-45.
8. Hsieh TC, Chow SC, Liu JP, Hsiao CF, Chi E (2010) Statistical test for evaluation of biosimilarity in variability of follow-on biologics. J Biopharm Stat 20: 75-89.